



## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **09281997 A**(43) Date of publication of application: **31.10.97**

(51) Int. Cl.

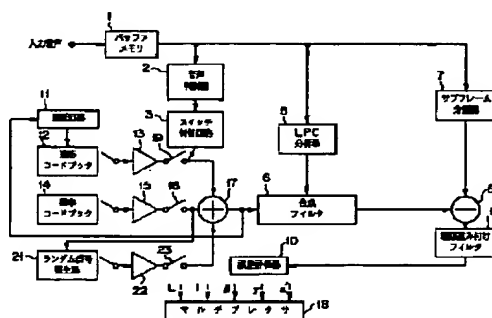
**G10L 9/14**  
**H03M 7/30**
(21) Application number: **08091178**(71) Applicant: **OLYMPUS OPTICAL CO LTD**(22) Date of filing: **12.04.96**(72) Inventor: **TAKAHASHI HIDEYUKI**(54) **VOICE CODING DEVICE**

## (57) Abstract:

**PROBLEM TO BE SOLVED:** To obtain a natural reproduced sound even in the case of coding a non-voice signal by setting the gain of a random signal expressing a sound source based on the gain at the time when a non-voice signal being an input signal at that time is coded by a voice sound source estimating part.

**SOLUTION:** When an input signal is discriminated to be a non-voice by a voice discriminator 2, since a switch 3 opens a switch 19, an adaptive code book 12 does not operate and the random signal made to be generated from a random signal generator 21 is used as a sound source signal. The gain of this random signal estimates a sound source signal by regarding the input signal at that time provisionally as a voice and by performing retrieval by a provability code look 14 temporally. Then, a random signal generator 21 calculates the energy of the estimated sound source and adjusts the gain of a random signal to be generated so that the energy of the random signal to be generated and the energy of the sound source become equal.

COPYRIGHT: (C)1997,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-281997

(43) 公開日 平成9年(1997)10月31日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 9/14			G 1 0 L 9/14	G
				C
				J
H 0 3 M 7/30		9382-5K	H 0 3 M 7/30	B

審査請求 未請求 請求項の数 2 O L (全 6 頁)

(21) 出願番号 特願平8-91178

(22) 出願日 平成8年(1996)4月12日

(71) 出願人 000000376

オリンパス光学工業株式会社

東京都渋谷区幡ヶ谷2丁目43番2号

(72) 発明者 ▲高▼橋 秀享

東京都渋谷区幡ヶ谷2丁目43番2号 オリ

ンパス光学工業株式会社内

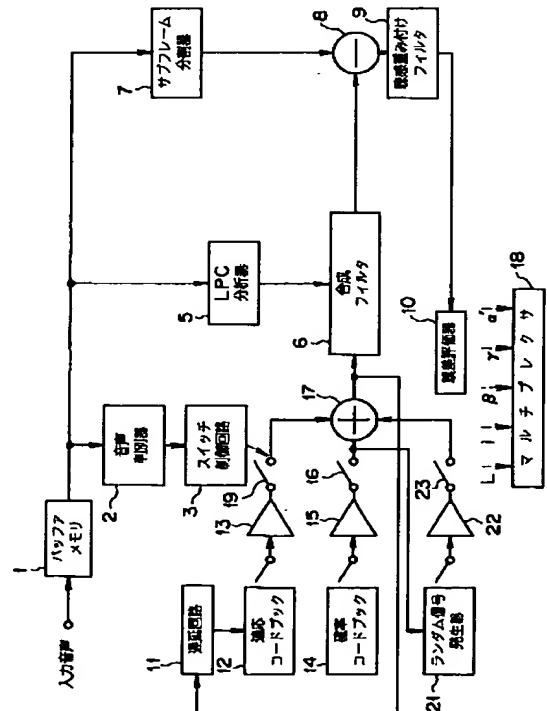
(74) 代理人 弁理士 伊藤 進

## (54) 【発明の名称】 音声符号化装置

## (57) 【要約】

【課題】 音声と非音声の境界が不自然でない再生音を得ることができる音声符号化装置を提供する。

【解決手段】 入力信号が音声信号か非音声信号かを判別する音声判別器2と、音声信号符号化用の遅延回路11、適応コードブック12、確率コードブック14等の音声音源推定部と、非音声信号符号化用のランダム信号発生器21等の非音声音源推定部と、上記音声判別器2による判別結果に基づいて上記音声音源推定部と非音声音源推定部の何れの動作を行うかを選択するスイッチ制御回路3とを備え、上記非音声音源推定部により符号化するデータは、線形予測パラメータと、音源信号を表すランダム信号と、このランダム信号のゲインを含み、該ランダム信号のゲインは、そのときの入力信号である非音声信号を上記音声音源推定部により符号化したときのゲインを所定の割合で抑圧した値に基づいて設定されるものである音声符号化装置。



## 【特許請求の範囲】

【請求項1】 入力信号が音声信号または非音声信号の何れであるかを判別する音声判別手段と、音声信号を符号化するための音声音源推定部と非音声信号を符号化するための非音声音源推定部とを有してなる線形予測符号化手段と、

上記音声判別手段による判別結果に基づいて、上記線形予測符号化手段により音声信号と非音声信号の何れの符号化を行うかを選択する符号化選択手段と、を備え、

上記非音声音源推定部により符号化するデータは、線形予測パラメータと、音源信号を表すランダム信号と、このランダム信号のゲインとを含み、該ランダム信号のゲインは、そのときの入力信号である非音声信号を上記音声音源推定部により符号化したときのゲインに基づいて設定されるものであることを特徴とする音声符号化装置。

【請求項2】 上記ランダム信号のゲインは、上記入力信号である非音声信号を上記音声音源推定部により符号化したときのゲインを所定の割合で抑圧した値に基づいて設定されるものであることを特徴とする請求項1に記載の音声符号化装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、音声符号化装置、より詳しくは、音声信号をデジタル情報圧縮して記録または伝送する音声符号化装置に関する。

## 【0002】

【従来の技術】音声信号を効率良く圧縮するために広く用いられている手段として、音声信号を、スペクトル包絡を表す線形予測パラメータと、線形予測残差信号に対応する音源パラメータとを用いて符号化する方式がある。このような線形予測の手段を用いた音声符号化方式は、少ない伝送容量で比較的高品質な合成音声を得られることから、最近のハードウェア技術の進歩と相まって様々な応用方式が盛んに研究され、開発されている。

【0003】また、さらに効率良く圧縮する技術として、音声の状態に応じてビット配分を変化させる可変レート符号化や高能率な音声符号化と非音声圧縮機能を組み合わせる方式が知られており、例えば特公平2-35996号公報に、その一例が記載されている。

## 【0004】

【発明が解決しようとする課題】しかしながら、上記特公平2-35996号公報に記載のものでは、音声を符号化するときと非音声を符号化するときとで、その符号化方法が大きく異なるために、特に音声区間と非音声区間との境界において、再生音が非常に不自然なものになってしまう難点がある。

【0005】本発明は上記事情に鑑みてなされたものであり、非音声信号を符号化する場合にも自然な再生音を

得ることができる音声符号化装置を提供することを目的としている。

## 【0006】

【課題を解決するための手段】上記の目的を達成するために、本発明の第1の音声符号化装置は、入力信号が音声信号または非音声信号の何れであるかを判別する音声判別手段と、音声信号を符号化するための音声音源推定部と非音声信号を符号化するための非音声音源推定部とを有してなる線形予測符号化手段と、上記音声判別手段による判別結果に基づいて上記線形予測符号化手段により音声信号と非音声信号の何れの符号化を行うかを選択する符号化選択手段とを備え、上記非音声音源推定部により符号化するデータは線形予測パラメータと音源信号を表すランダム信号とこのランダム信号のゲインとを含み、該ランダム信号のゲインはそのときの入力信号である非音声信号を上記音声音源推定部により符号化したときのゲインに基づいて設定されるものである。

【0007】また、本発明の第2の音声符号化装置は、上記第1の音声符号化装置において、上記ランダム信号のゲインが、上記入力信号である非音声信号を上記音声音源推定部により符号化したときのゲインを所定の割合で抑圧した値に基づいて設定されるものである。

【0008】従って、本発明の第1の音声符号化装置は、音声判別手段が入力信号が音声信号または非音声信号の何れであるかを判別し、線形予測符号化手段の音声音源推定部が音声信号を符号化するとともに非音声音源推定部が非音声信号を符号化し、符号化選択手段が上記音声判別手段による判別結果に基づいて上記線形予測符号化手段により音声信号と非音声信号の何れの符号化を行うかを選択し、上記非音声音源推定部により符号化するデータは線形予測パラメータと音源信号を表すランダム信号とこのランダム信号のゲインとを含み、該ランダム信号のゲインはそのときの入力信号である非音声信号を上記音声音源推定部により符号化したときのゲインに基づいて設定される。

【0009】また、本発明の第2の音声符号化装置は、上記ランダム信号のゲインが、上記入力信号である非音声信号を上記音声音源推定部により符号化したときのゲインを所定の割合で抑圧した値に基づいて設定される。

## 【0010】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態を説明する。図1から図3は本発明の一実施形態を示したものであり、図1は音声符号化装置の構成を示すブロック図である。

【0011】本実施形態の音声符号化装置は、コード駆動線形予測符号化(CELP)方式の音声符号化装置に非音声圧縮機能を付加したものである。

【0012】図1に示すように、入力端子に接続されたバッファメモリ1の出力端は3つに分岐されていて、第1の出力端はサブフレーム分割器7を介して減算器8に

接続され、第2の出力端は線形予測分析手段たるLPC分析器5の入力端に接続され、第3の出力端は音声判別手段たる音声判別器2を介して符号化選択手段たるスイッチ制御回路3に接続されている。

【0013】このスイッチ制御回路3は、上記音声判別器2による判別結果に応じて後述するスイッチ19の開閉動作を制御するものである。すなわち、音声判別器2において入力信号が音声であると判定されればスイッチ19を閉じて後述する適応コードブック12を有効にし、非音声であると判定されればスイッチ19を開いて適応コードブック12を無効にするようになっている。

【0014】上記LPC分析器5は、合成フィルタ6に接続されていて、この合成フィルタ6には、音声音源推定部を構成する適応コードブック12と同音声音源推定部を構成する確率コードブック14とを用いて生成される信号が入力されるようになっている。

【0015】すなわち、上記適応コードブック12は、乗算器13とスイッチ19とを介して加算器17の第1入力端子に接続されており、また、確率コードブック14は、乗算器15とスイッチ16とを介して上記加算器17の第2入力端子に接続されている。

【0016】この加算器17の出力端子は、合成フィルタ6を介して上記減算器8の入力端子に接続されるとともに、音声音源推定部を構成する遅延回路11を介して上記適応コードブック12に接続されている。

【0017】上記確率コードブック14の出力は、乗算器15とスイッチ16とを介して非音声音源推定部たるランダム信号発生器21にも出力されるようになっている。このランダム信号発生器21の出力は、乗算器22およびスイッチ23を介して上記加算器17の第3入力端子に接続されている。

【0018】上記合成フィルタ6の出力端は、サブフレーム分割器7が接続された減算器8および聴感重み付けフィルタ9を介して誤差評価器10の入力端子に接続されている。この誤差評価器10による評価結果は、マルチプレクサ18に出力されるようになっている。

【0019】上述のような音声符号化装置において、線形予測符号化手段は、上記遅延回路11、適応コードブック12、確率コードブック14、ランダム信号発生器21、乗算器13、15、22、スイッチ16、19、23、加算器17等を含んで構成されている。

【0020】次に、図2は上記音声判別器2のより詳細な構成を示すブロック図である。

【0021】この音声判別器2に入力された上記バッファメモリ1の出力信号は、2つに分岐されて一方がフレームエネルギー分析回路2aに、他方が初期フレームエネルギー分析回路2bに入力されるようになっている。

【0022】上記フレームエネルギー分析回路2aは加算器2cの+端子となっている第1入力端子に、上記初期フレームエネルギー分析回路2bは該加算器2cの-

端子となっている第2入力端子にそれぞれ接続されているとともに、さらに、初期フレームエネルギー分析回路2bは、閾値決定回路2dにも接続されている。

【0023】そして、上記加算器2cの出力端子と上記閾値決定回路2dの出力端子は、共に判別回路2eに接続されていて、この判別回路2eの出力が上記スイッチ制御回路3に入力されるようになっている。

【0024】次に、上記図1および図2に示したような構成における信号の流れを説明する。

10 【0025】入力端子から例えば8kHz（すなわち、1サンプル当たり1/8ms）でサンプリングされた原音信号を入力して、予め定められたフレーム間隔（例えば20ms、すなわち160サンプル）の音信号をバッファメモリ1に格納する。

【0026】バッファメモリ1は、入力信号をフレーム単位でサブフレーム分割器7とLPC分析器5と音声判別器2とに送出する。

20 【0027】この音声判別器2は、フレームの入力信号が音声か非音声かを、例えば以下に説明するような方法で判別する。

【0028】上記図2に示したような構成の音声判別器2において、フレームエネルギー分析回路2aは、入力されたフレーム入力信号のフレームエネルギー $E_f$ を次に示すような数式により算出する。

【0029】

【数1】

$$E_f = 10 \log_{10} \left\{ 1 + \sum_{n=0}^{N-1} s^2(n) \right\}$$

30 ここに、 $s(n)$ はサンプル $n$ における入力信号、 $N$ はフレーム長をそれぞれ示している。

【0030】また、上記初期フレームエネルギー分析回路2bは、符号化を開始したときのフレームエネルギー $E_b$ を上記数式1と同様の数式を用いて算出する。

40 【0031】上記閾値決定回路2dは、背景雑音エネルギーの大きさに応じて閾値を決定する。例えば、図3に示すように、背景雑音エネルギーがdB単位で増加するに従って、閾値をdB単位で減少させる関係により、閾値を決定する。そして、その結果を判別回路2eに送出する。

【0032】加算器2cでは、フレームエネルギー $E_f$ を正として入力するとともに、初期フレームエネルギー $E_b$ を負として入力してこれらを加算することにより、フレームエネルギー $E_f$ から初期フレームエネルギー $E_b$ を減算し、その減算結果を判別回路2eに送出する。

【0033】そして、判別回路2eは、入力された減算結果と閾値を比較して、減算結果が閾値より大きければフレーム入力信号は音声区間であると判別し、そうでなければ非音声区間であると判別する。

50 【0034】図1に戻って、サブフレーム分割器7は、

フレームの入力信号を予め定められたサブフレーム間隔（例えば5ms、つまり40サンプル）に分割する。すなわち、1フレームの入力信号から、第1サブフレームから第4サブフレームまでの4つのサブフレーム信号が作成される。

【0035】LPC分析器5は、入力信号に対して線形予測分析（LPC分析）を行って、スペクトル特性を表すスペクトルパラメータたる線形予測パラメータ $\alpha$ を抽出し、合成フィルタ6およびマルチプレクサ18に送出する。

【0036】また、適応コードブックの遅れ $L$ 、ゲイン $\beta$ 、確率コードブックのインデックス $i$ 、ゲイン $\gamma$ は、次に説明するような手段により決定される。

【0037】まず、適応コードブックの遅延 $L$ とゲイン $\beta$ は、以下の処理によって決定される。

【0038】遅延回路11において、先行サブフレームにおける合成フィルタ6の入力信号すなわち駆動音源信号に、ピッチ周期に相当する遅延を与えて適応コードベクトルとして作成する。

【0039】例えば、想定するピッチ周期を40～167サンプルとすると、40～167サンプル遅れの128種類の信号が適応コードベクトルとして作成され、適応コードブック12に格納される。

【0040】このときスイッチ16は開いた状態となっていて、各適応コードベクトルは乗算器13でゲイン値を可変して乗じた後に、加算器17を通過してそのまま合成フィルタ6に入力される。

【0041】この合成フィルタ6は、線形予測パラメータ $\alpha'$ を用いて合成処理を行い、合成ベクトルを減算器8に送出する。この減算器8は、原音声ベクトルと合成ベクトルとの減算を行うことにより誤差ベクトルを生成し、得られた誤差ベクトルを聴感重み付けフィルタ9に送出する。

【0042】この聴感重み付けフィルタ9は、誤差ベクトルに対して聴感特性を考慮した重み付け処理を行い、誤差評価器10に送出する。

【0043】誤差評価器10は、誤差ベクトルの2乗平均を計算し、その2乗平均値が最小となる適応コードベクトルを検索して、その遅れ $L$ とゲイン $\beta$ をマルチプレクサ18に送出する。このようにして、適応コードブック12の遅延 $L$ とゲイン $\beta$ が決定される。

【0044】続いて、確率コードブックのインデックス $i$ とゲイン $\gamma$ は、以下の処理によって決定される。

【0045】確率コードブック14は、サブフレーム長に対応する次元数（すなわち、上述の例では40次元）の確率コードベクトルが、例えば512種類予め格納されており、各々にインデックスが付与されている。なお、このときにはスイッチ16は閉じた状態となっている。

【0046】まず、上記処理によって決定された最適な

適応コードベクトルを、乗算器13で最適ゲイン $\beta$ を乗じた後に、加算器17に送出する。

【0047】次に、各確率コードベクトルを乗算器15でゲイン値を可変して乗じた後に、加算器17に入力する。加算器17は上記最適ゲイン $\beta$ を乗じた最適な適応コードベクトルと各確率コードベクトルの加算を行い、その結果が合成フィルタ6に入力される。

【0048】この後の処理は、上記適応コードブックパラメータの決定処理と同様に行われる。すなわち、合成フィルタ6は線形予測パラメータ $\alpha'$ を用いて合成処理を行い、合成ベクトルを減算器8に送出する。

【0049】減算器8は原音声ベクトルと合成ベクトルとの減算を行うことにより誤差ベクトルを生成し、得られた誤差ベクトルを聴感重み付けフィルタ9に送出する。

【0050】聴感重み付けフィルタ9は、誤差ベクトルに対して聴感特性を考慮した重み付け処理を行い、誤差評価器10に送出する。

【0051】誤差評価器10は、誤差ベクトルの2乗平均を計算して、その2乗平均値が最小となる確率コードベクトルを検索して、そのインデックス $i$ とゲイン $\gamma$ をマルチプレクサ18に送出する。このようにして、確率コードブック14のインデックス $i$ とゲイン $\gamma$ が決定される。

【0052】上記マルチプレクサ18は、量子化された線形予測パラメータ $\alpha'$ 、適応コードブックの遅れ $L$ 、ゲイン $\beta$ 、確率コードブックのインデックス $i$ 、ゲイン $\gamma$ の各々をマルチプレクスして伝送するものである。

【0053】次に、非音声であると判定されたときの音源信号のゲインの設定方法について説明する。

【0054】音声判別器2により入力信号が非音声であると判定されると、まず、音声時と同様にしてLPC分析が行われる。これにより、非音声においてもLPC情報が送出されるために入力信号の特徴をある程度保つことができる。

【0055】また、非音声であると判定された場合には、上述したように、スイッチ制御回路3がスイッチ19を開くために適応コードブック12は動作せず、音源信号としてはランダム信号発生器21から発生させたランダム信号を用いている。

【0056】上記ランダム信号のゲインは、そのときの入力信号が仮に音声であるとみなして、一旦、確率コードブック14による探索を行い、音源信号を推定する。

【0057】そして、ランダム信号発生器21は、推定された音源信号のエネルギーを計算し、発生するランダム信号のエネルギーと音源信号のエネルギーとが等しくなるように、発生するランダム信号のゲインを調節する。

【0058】さらに、そのランダム信号のゲイン値は、そのときの入力信号に対して仮に音声とみなしたときに

推定される音源信号のエネルギーを、所定の割合、例えば $1/2$ に抑圧した値に設定する。

【0059】このように音源信号のゲインを、仮に音声とみなしたときに推定される音源信号のエネルギーに基づいて設定されるために、音声区間と非音声区間の境界で発生する不自然さを軽減することができる。

【0060】また、入力信号に対して仮に音声とみなしたときに推定される音源信号のエネルギーを所定の割合に抑圧した値に設定するために、非音声時に耳につき易い背景ノイズを軽減することができる。

【0061】なお、音声判別器における音声判別方法は、上述した手段に限るものではないことはいうまでもない。

【0062】また、上記実施形態においては、コード駆動線形予測符号化装置を一例として取り上げて説明したが、線形予測パラメータと、線形予測残差信号に相当する駆動音源信号のパラメータとで表現する符号化装置であれば、当然にして、何れのものにも適用することが可能である。

【0063】このような実施形態によれば、非音声時の音源信号のゲインを、仮に音声とみなしたときに推定される音源信号のエネルギーに基づいて設定するために、音声区間と非音声区間の境界で発生する不自然さを軽減することができ、非音声圧縮機能を用いても自然な再生音を得ることができる音声符号化装置となる。

【0064】

【発明の効果】以上説明したように請求項1に記載の発明 \*

\* 明によれば、音源信号を表すランダム信号のゲインを、そのときの入力信号である非音声信号を音声音源推定部により符号化したときのゲインに基づいて設定するために、非音声信号を符号化する場合にも自然な再生音を得ることができる。

【0065】また、請求項2に記載の発明によれば、請求項1に記載の発明と同様の効果を奏するとともに、非音声であるときの背景雑音を抑制することができる。

【図面の簡単な説明】

10 【図1】本発明の一実施形態の音声符号化装置の構成を示すブロック図。

【図2】上記実施形態の音声判別器のより詳細な構成を示すブロック図。

【図3】上記実施形態において、音声判別器の閾値決定回路により決定される閾値と背景雑音エネルギーとの関係の一例を示す線図。

【符号の説明】

2…音声判別器（音声判別手段）

3…スイッチ制御回路（符号化選択手段）

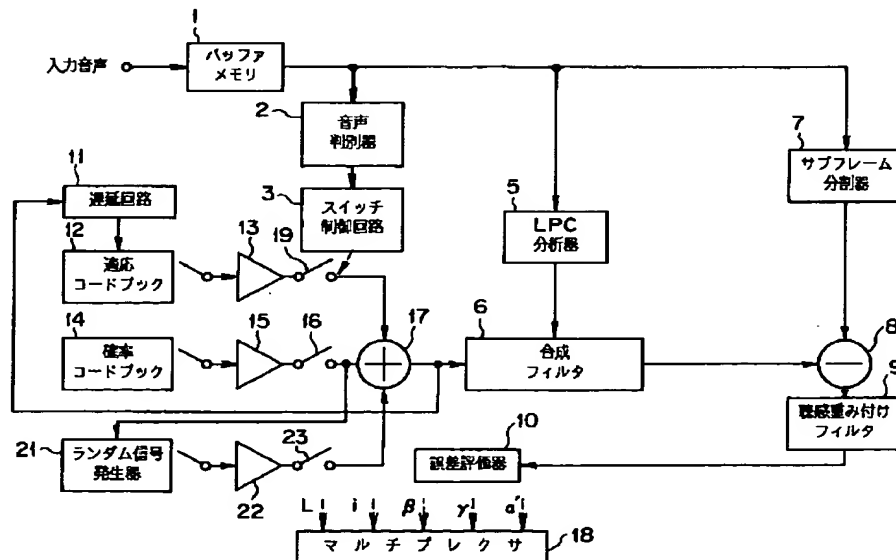
20 11…遅延回路（音声音源推定部，線形予測符号化手段の一部）

12…適応コードブック（音声音源推定部，線形予測符号化手段の一部）

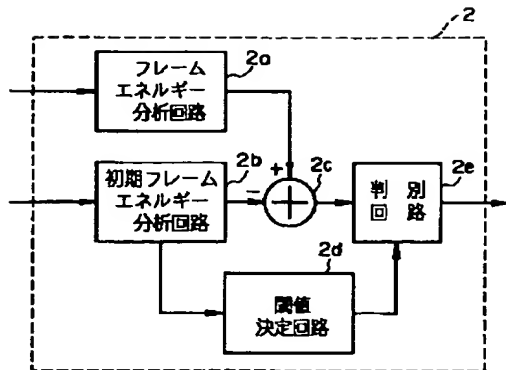
14…確率コードブック（音声音源推定部，線形予測符号化手段の一部）

21…ランダム信号発生器（非音声音源推定部，線形予測符号化手段の一部）

【図1】



【図 2】



【図 3】

